Isolation and molecular identification of sunshine virus, a novel paramyxovirus found in australian snakes

Timothy H. Hyndman, Rachel E. Marschang, James F.X. Wellehan, Philip K. Nicholls

PII: S1567-1348(12)00168-2
DOI: http://dx.doi.org/10.1016/j.meegid.2012.04.022
Reference: MEEGID 1292

To appear in: *Infection, Genetics and Evolution*

Please cite this article as: Hyndman, T.H., Marschang, R.E., Wellehan, J.F.X., Nicholls, P.K., Isolation and molecular identification of sunshine virus, a novel paramyxovirus found in australian snakes, *Infection, Genetics and Evolution* (2012), doi: http://dx.doi.org/10.1016/j.meegid.2012.04.022

Title Page

ISOLATION AND MOLECULAR IDENTIFICATION OF SUNSHINE VIRUS, A NOVEL PARAMYXOVIRUS FOUND IN AUSTRALIAN SNAKES

Timothy H Hyndman[a]*, Rachel E Marschang[b], James F.X. Wellehan Jr.[c] & Philip K Nicholls[a]

[a] School of Veterinary and Biomedical Sciences, Murdoch University, Murdoch, Western Australia, 6150, Australia. t.hyndman@murdoch.edu.au, p.nicholls@murdoch.edu.au

[b] Institut für Umwelt und Tierhygiene, University of Hohenheim, Garbenstr. 30, 70599 Stuttgart, Germany. rachel.marschang@googlemail.com

[c] Department of Small Animal Clinical Sciences, University of Florida College of Veterinary Medicine, Gainesville, FL, 32610 USA. wellehanj@ufl.edu

*Corresponding author: phone: +61 08 9360 7348, fax: +61 08 9310 4144

Abstract

This paper describes the isolation and molecular identification of a novel paramyxovirus found during an investigation of an outbreak of neurorespiratory disease in a collection of Australian pythons. Using Illumina® high-throughput sequencing, a 17,187 nucleotide sequence was assembled from RNA extracts from infected viper heart cells (VH2) displaying widespread cytopathic effects in the form of multinucleate giant cells. The sequence appears to contain all the coding regions of the genome, including the following predicted paramyxoviral open reading frames (ORFs): 3' – Nucleocapsid (N) – putative Phosphoprotein (P) – Matrix (M) – Fusion (F) – putative attachment protein – Polymerase (L) – 5'. There is also a 540 nucleotide ORF between the N and putative P genes that may be an additional coding region. Phylogenetic analyses of the complete N, M, F and L genes support the clustering of this virus within the family *Paramyxoviridae* but outside both of the current subfamilies: *Paramyxovirinae* and *Pneumovirinae*. We propose to name this new virus, Sunshine virus, after the geographic origin of the first isolate – the Sunshine Coast of Queensland, Australia.

Keywords

Reptiles, Snakes, *Paramyxoviridae*, High-throughput nucleotide sequencing, Phylogeny

1. Introduction

Members of the family *Paramyxoviridae* are enveloped single-stranded negative-sense RNA viruses and are currently divided into two subfamilies, *Pneumovirinae* and *Paramyxovirinae*, which contain seven and two genera respectively (ICTV, 2012). Previous to this report, all known paramyxoviruses utilising squamate hosts (snakes and lizards) clustered within the genus *Ferlavirus* (Franke et al., 2001; Marschang et al., 2009; Papp et al., 2010).

An outbreak of neurorespiratory disease in a Swiss serpentarium in 1972 formed the basis for the first description of the isolation of a paramyxovirus from a snake (Folsch and Leloup, 1976). Since this time, paramyxoviruses have been isolated from similarly-affected snakes from other regions of Europe (Ahne et al., 1987; Blahak, 1995; Franke et al., 2001; Manvell et al., 2000), USA (Jacobson et al., 1981; Jacobson et al., 1980; Potgieter et al., 1987; Richter et al., 1996) and Brazil (Kolesnikovas et al., 2006; Nogueira et al., 2002). Koch's postulates have been fulfilled in five Aruba Island rattlesnakes (*Crotalus unicolor*) to imply a causative association between ferlaviral infection and disease (Jacobson et al., 1997). Besides snakes, ferlaviruses have been associated with disease in tortoises (Marschang et al., 2009; Zangger et al., 1991) and lizards (Jacobson et al., 2001; Marschang et al., 2009). This paper describes the isolation and initial studies of a novel paramyxovirus discovered in a private Australian snake collection that was associated with significant morbidity and some mortalities. The novel paramyxovirus described in this paper is distantly related to the *Ferlavirus* genus and we suggest this new virus be named *Sunshine virus*, after the origin of the first isolate – the Sunshine Coast of Queensland, Australia.

2. Materials and Methods

2.1 Sample collection

In 2008, a private breeder of birds and reptiles from Queensland, Australia acquired seven jungle carpet pythons (*Morelia spilota cheynei*) as a breeding loan from another private keeper. The snakes were to be added to a collection of 70 Australian pythons (*Antaresia* sp., *Morelia spilota* ssp. and *Aspidites* sp.). Following an outbreak of neurorespiratory disease, the entire collection was humanely euthanased. Samples were collected from these animals and then submitted to the primary author (THH) for further investigation. In total, samples from 17 livers, kidneys and lungs, 16 brains and 13 serum samples were collected from 17 snakes. Snakes were selected for sample collection based on clinical signs and/or which snakes they had been in direct contact with. 12 of the 17 snakes were symptomatic while the remaining five snakes were in-contact but asymptomatic. Half of each organ was submitted frozen and the other half was submitted fixed in formalin for histopathological examination.

2.2 Virus Isolation

Liver and kidney samples from each animal were pooled but lung and brain samples were tested individually. In total, 50 samples (17 pooled liver-kidney, 17 lungs and 16 brains) were processed for virus isolation. Virus isolation was not attempted on any serum sample. Approximately one cubic centimetre ($1cm^3$) pieces of fresh frozen organ were individually placed into aliquots of 2mL of virus isolation media which contained minimum essential Eagle's medium with Earle's salts (MEM, Sigma, cat. no. M5650) supplemented with 5% (v/v) foetal bovine serum (FBS, GIBCO, cat. no. 10100-147), 2 x enrofloxacin (25µg/mL using Baytril® 2.5% Oral Solution, Bayer), 2 x amphotericin B (5µg/mL, GIBCO, cat. no. 15290-018), 5 x penicillin G/streptomycin (50IU/mL and 0.5mg/mL respectively, Sigma, cat. no. P4333-20ML) and 1 x L-glutamine (2mM, GIBCO, cat. no. 25030-149). Samples were then aseptically and finely diced using sterile scissors. Samples were then vigorously vortexed and clarified (4,000g x 10min @ 4˚C).

Viper heart cells (VH2, ATCC CCL-140) were grown at 30˚C and 5% CO_2. At 80%-100% confluency, the cell culture medium was removed, the cells were rinsed with 1 x phosphate-buffered saline (PBS) and supernatant from the clarified tissue suspension was added to the cells. Flasks and plates were left to incubate at room temperature for 1 hour. The tissue suspension supernatant was removed, cells were rinsed several times with 1 x PBS and then virus isolation media was added. Cells were observed daily for cytopathic effects (CPE). Seven days after inoculation, cells were frozen and then left to thaw at room temperature. Medium was clarified as above and supernatant was then used to replace the maintenance media of 75%-80% confluent VH2 cells. Wells were left at room temperature for one hour and then maintenance media was added. For some virus isolation attempts, this freeze-thawing passage was repeated once more. Viral titre was determined using the Reed-Muench method as previously described (Mahy and Kangro, 1996).

2.3 Polymerase chain reaction and High-throughput sequencing

For the purposes of polymerase chain reaction (PCR), aliquots of unprocessed media and frozen-thawed, clarified cell lysate were taken from infected flasks. Nucleic acid was extracted using the MagMAX™ Viral RNA Isolation Kit (Ambion, Austin, Texas, cat. no. AM1939) according to the manufacturer's protocols. RNA was reverse transcribed into cDNA using Superscript® III reverse transcriptase (Invitrogen, Mulgrave, Victoria, cat. no. 18080044) and either random hexamers or

gene-specific primers (data not shown). Primers that were tested on the extracted nucleic acid are listed in the Appendix in Table A.1. For some of these PCRs, the stringency was lowered by decreasing the annealing temperature (data not shown).

For Illumina® high-throughput sequencing, RNA was extracted from a 25cm^2 flask of VH2 cells, infected with Sunshine virus that was displaying extensive CPE. Medium was removed and the monolayer was rinsed with PBS. 1mL of Trizol® LS (Invitrogen, Mulgrave, Victoria, cat. no. 10296010) was added to the flask, pipetted thoroughly and then transferred to a new tube. 333µL of chloroform was added, the tube was vortexed and then left at room temperature for five minutes. The contents of the tube were then transferred to a phase lock gel heavy separator tube and centrifuged at 12,000g for 15 minutes at 4˚C. The volume of clear supernatant was mixed with 100% ethanol to an ethanol concentration of 33% (v/v). This was added to an RNeasy® spin column (Qiagen, Doncaster, Victoria, cat. no. 74104) and centrifuged at 8,000g for one minute. Flow through was discarded and the remaining washes were performed in accordance with the manufacturer's instructions. Ribosomal RNA (rRNA) was removed from total RNA using the Ribominus™ Eukaryote Kit (Invitrogen, Mulgrave, Victoria, cat. no. A10837-02) and the rRNA-depleted RNA was sent to Fasteris (Geneva, Switzerland) for further processing. Fasteris then performed the following steps: zinc breakage of RNA, cDNA synthesis, ends repair, adaptor ligation, gel purification, PCR amplification, Illumina® sequencing and *de novo* assembly.

5,818 unique contigs were assembled. The Basic Local Alignment Search Tool (BLASTN; http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to try to determine the identity of these unique sequences. Batch searches were performed and the results downloaded. Results were then searched for the word "virus". From the 5,818 unique contigs, 326 had BLAST hits for the word "virus". The viral family was then determined for each BLAST hit and viruses not known to occur in vertebrates were excluded. This excluded host DNA (e.g. virus receptors) and viruses of algae, fungi, invertebrates, plant and protozoa. This left 212 unique contigs representing 25 classified families. Next, ORF Finder (http://www.ncbi.nlm.nih.gov/projects/gorf/) was used to identify open reading frames (ORF) and then search the protein database of GenBank (National Center for Biotechnology Information,Maryland) for similarities. From the 212 unique contigs that were analysed in this way, only one (an 11,709 nucleotide contig) showed strong similarity to a virus. This sequence had one ORF that was similar to the fusion glycoprotein superfamily and a second ORF that was similar to paramyxovirus RNA-dependent RNA-polymerase.

The raw sequencing data was then reassembled using CLC Genomics Workbench® software (CLC Bio©) to look for additional sequence information belonging to the putative paramyxoviral genome and a 17,187 nucleotide sequence was assembled: 5,478 nucleotides longer than the 11,709 nucleotide contig assembled by Fasteris®.

Using the putative paramyxoviral sequence information that had been generated, non-degenerate primers were then designed (Table 1) based on the RNA-dependent RNA-polymerase gene. These primers would then be used in PCRs to look for Sunshine virus in other samples (manuscript submitted). To determine that this new virus was not endogenous to the VH2 cell line, all three primer pairs were used on infected and uninfected cells.

RNA was extracted from infected and uninfected VH2 cells as above. cDNA synthesis was performed using random hexamers and Superscript® III under the cycling conditions: 25˚C for 5 minutes, 45˚C

for 45 minutes and then the reverse transcriptase was heat inactivated at 70˚C for 15 minutes. For PCR amplification, each primer (1µM final concentration) was added to 1µL of cDNA and then Platinum® PCR Supermix (Invitrogen, Mulgrave, Victoria, cat. no. 11306016) was used to bring the final reaction volume to 20µL. Cycling conditions were as follows: 94˚C x 2min, 40 x (94˚C x 20s, 45˚C x 45s, 72˚C x 30s). 45˚C was used as the annealing temperature for all four Sunshine virus primers. PCR products were visualised using agarose gel electrophoresis and then sequenced with an AB3730xl-capillary sequencer (Applied Biosystems, California). Sequence data was aligned with the 17,187 nucleotide contig obtained from high-throughput sequencing.

2.5 Phylogenetic analysis

Open reading frames (ORFs) were identified using a web-based translator (http://www.vivo.colostate.edu/molkit/translate/index.html). Large ORFs were identified by sequence homology to known paramyxoviruses in GenBank (National Center for Biotechnology Information,Maryland), EMBL (Cambridge, United Kingdom), and Data Bank of Japan (Mishima, Japan) databases using BLASTP (Altschul et al., 1997), and by location in the genome.

The predicted homologous amino acid sequences of mononegaviral proteins were aligned using MAFFT (Katoh and Toh, 2008). Bayesian analyses of each alignment were performed using MrBayes 3.1 (Ronquist and Huelsenbeck, 2003) on the CIPRES server (Miller et al., 2010), with gamma distributed rate variation and a proportion of invariant sites, and amino acid substitution model jumping. Four chains were run and statistical convergence was assessed by looking at the standard deviation of split frequencies as well as potential scale reduction factors of parameters. The first 10% of 1,000,000 iterations were discarded as a burn in, based on examination of trends of the log probability vs. generation. Two independent Bayesian analyses were run to avoid entrapment on local optima.

Maximum likelihood (ML) analyses of each alignment were performed using RAxML on the CIPRES server (Stamatakis et al., 2008) , with gamma distributed rate variation and a proportion of invariant sites. The amino acid substitution model with the highest posterior probability in the Bayesian analysis was selected. Nyamanini virus (GenBank accession number NC012703) was used as the outgroup for the N and L genes, Bornavirus (NC_001607) was used as the outgroup for the M gene, and Atlantic salmon paramyxovirus (EF646380) was designated as the outgroup for the F gene. Bootstrap analysis was used to test the strength of the tree topology (200 re-samplings) (Felsenstein, 1985). Numbers of bootstrap replicates were determined using the stopping criteria by Pattengale *et al.* (2010).

To test for paraphyly of the *Paramyxoviridae*, likelihood ratio testing was conducted (Shimodaira and Hasegawa, 1999). Trees of the nucleoprotein were constrained to paramyxoviral monophyly and compared by the Shimodaira-Hasegawa test to the best unconstrained tree identified in RAxML.

3. Results

3.1 Virus Isolation

After the inoculation of 50 clarified tissue suspensions from 12 clinically affected and five in-contact asymptomatic snakes, VH2 cells were examined daily for cytopathic effects (CPE). During the first seven days of observation, no CPE was detected. However, CPE was seen within one to two days

after the first blind passage with brain, lung and combined kidney-liver from a black-headed python (*Aspidites melanocephalus*), and brain from a jungle carpet python (*Morelia spilota cheynei*). For all four isolates, CPE was characterised by extensive formation of multinucleate giant cells (Figure 1). Variable amounts of cell lysis were seen. Cell rounding, cell detachment, perinuclear granulation and cytoplasmic vacuolation were not significant hallmarks of this virus's CPE. It was found that more widespread CPE could be obtained by trypsinising and passaging the cells (one to three split) at the first sign of CPE. The isolate from the black-headed python lung suspension was selected for all further testing. The tissue culture infectious dose ($TCID_{50}$) of this isolate was calculated as $10^{2.75}$ mL^{-1}. In total, this isolate was serially passaged four times without any observable changes to CPE. Additional passages were not attempted.

3.2 Polymerase chain reaction and High-throughput sequencing

In our hands, eight consensus PCRs capable of detecting a broad range of paramyxoviruses, were unable to detect Sunshine virus (Table A.1).

Using Illumina® sequencing, 10,544,936 reads (each 38 nucleotides long) were generated from the rRNA-depleted total RNA. A 17,187 nucleotide contig was assembled from 292,587 reads (2.77% of total reads). The mean coverage of this 17,187 nucleotide contig was 648 reads with a standard deviation of 605 reads. There was a minimum of one read and a maximum of 4,647 reads over the length of this 17,187 nucleotide contig. The first 21 nucleotides from the 3' end of this 17,187 nucleotide contig had 16 or fewer reads, while the last four nucleotides at the 5' end had four or fewer reads. It is therefore possible that the 17,187 nucleotide contig does not represent the entire genome of Sunshine virus. The remaining 17,162 nucleotides (of the 17,187 nucleotide contig) had 20 or more reads contributing to each nucleotide of this contig.

The 17,187 nucleotide sequence has been deposited into GenBank (ID: JN192445).

From all of the organ cultures showing CPE, all three PCR primer sets (S1-AS1, S2-AS2 and S2-AS1) yielded an amplicon, which when sequenced, showed 100% nucleotide identity to the relevant segment of the 17,187 nucleotide sequence. PCR results of the uninfected VH2 cells were negative.

3.3 Partial Genome Characterisation

Figure 2 summarises the features of the 17,187 nucleotide sequence. This sequence contains open reading frames (ORFs) clearly identified by strong homology as paramyxoviral nucleocapsid (N), matrix (M), fusion (F) and polymerase (L) ORFs. Between F and L, lies an ORF of 2,031 nucleotides. In all paramyxoviruses except the genus *Pneumovirus*, this is the location of the attachment gene (H, HN or G). Between N and M, lies a 540 nucleotide ORF and a 1,677 nucleotide ORF. In all paramyxoviruses, the phosphoprotein (P) is located between N and M, and when compared with other mononegaviral proteins on a BLASTP search, the top hit for the 1,677 nucleotide ORF is weak homology with the P of vesicular stomatitis virus, a rhabdovirus (e=0.08). The 540 nucleotide ORF lies between N and the putative P; only paramyxoviruses from the genus *Ferlavirus* (the other known paramyxoviruses utilising squamate hosts) are known to translate a protein (U) between N and P (Kurath et al., 2004; Marschang et al., 2009). However, these taxa are not closely related, there is no sequence homology, and therefore no reason to expect that these ORFs are orthologous.

Predicted intergenic regions of Sunshine virus were relatively large, ranging from 159 to 502 nucleotides. The conserved motifs of Sunshine virus that are consistent with paramyxoviral gene

transcription start and end signals are identified in Table 2. Transcriptional start and stop sites are conserved motifs in the *Paramyxoviridae*. In the genomic antisense, gene starts are typically uracil/cytosine-rich motifs eight nucleotides in length with a terminal uracil (Lamb and Parks, 2007). Gene ends are typically G/C/U rich motifs followed by 4-7 uracils (Lamb and Parks, 2007). A 3'-CUCUCuCU-5' motif (capital letters are completely conserved) was identified before every predicted ORF except the fusion gene, which may serve as a gene start site (Table 2). A CUCUCUCU motif was identified 87 nucleotides after the predicted start codon of the fusion gene. A 3'- RuUuaa$(U)_{4-8}$ motif (R=A or G, capital letters are completely conserved) was identified after every predicted ORF except the matrix gene, which may serve as a gene stop site (Table 2). A UAUUAAUUUUUUUU motif was identified 19 nucleotides after the predicted start codon of the fusion gene, which may serve as the transcription stop site for the matrix gene. Potential start and stop sites identified between the M and F genes diverged more from the predicted motifs. These sites may be functional, or the start codon of the F gene may actually be the second ATG of the predicted gene (which would eliminate the lysine-rich predicted first 57 amino acids), or M and F are transcribed as a bicistronic mRNA. Further study is needed to evaluate these possibilities. Bicistronic M/F mRNA has been seen in several of the *Paramyxovirinae*, and it is the most common site for readthrough transcription of many paramyxoviruses (Rassa and Parks, 1998; Yu et al., 1992).

In the N protein of all known members of the subfamily *Paramyxovirinae*, the most conserved motif is suggested to be responsible for N–N self-assembly, F-X_4-Y-X_3- -S- -A-M-G, where is any aromatic amino acid (Myers et al., 1997). The homologous region of the N protein of Sunshine virus, FAPAEYSNLYSFAIG, differs from this in the replacement of the methionine with an isoleucine.

In all characterized viruses in the subfamily *Paramyxovirinae*, but not *Pneumovirinae*, the P gene contains a conserved motif that is involved in RNA editing, so that other reading frames can be utilized (V/W/D proteins) (Kolakofsky et al., 2005). A homologous motif could not be identified in the putative P gene of Sunshine virus, further serving to distinguish it from the *Paramyxovirinae*.

3.4 Phylogenetic analysis

Bayesian phylogenetic analysis of the predicted N and M proteins found that the WAG model of amino acid substitution was most probable with a posterior probability of 1.000 (Whelan and Goldman, 2001). The predicted F protein analysis found the CpRev model to be most probable (posterior probability=1.000), and the predicted L protein analysis found the Blosum model to be most probable (posterior probability=1.000) (Adachi et al., 2000; Henikoff and Henikoff, 1992). Bayesian trees including posterior probabilities of clades are shown (Figures 3-6).

ML analysis was in agreement with the Bayesian analysis. Bootstrap values from ML analysis are shown on the Bayesian trees (Figures 3-6).

None of the analyses found that Sunshine virus clustered within either currently recognized subfamily of the paramyxoviruses. The predicted N and L protein analyses found that Sunshine virus clustered as a sister group with the *Paramyxoviridae* (Figure 3). The F protein analysis found support for genera and subfamilies, but homologous genes are not present in other members of the *Mononegavirales* to examine deeper level relationships. The M protein analysis, while supporting recognized genera as monophyletic, was not able to resolve deeper level relationships, including not forming a distinct monophyletic cluster from the *Avulavirus*/*Rubulavirus* and

Respirovirus/*Henipavirus*/*Morbillivirus* clades in the subfamily *Paramyxovirinae*. The nucleoprotein analysis found paraphyly of the family *Paramyxoviridae*, with support for a clade containing *Rhabdoviridae*/*Paramyxovirinae*/Sunshine virus and a clade containing *Filoviridae*/*Pneumovirinae*. While the nucleoprotein ML analysis found the same topology as Bayesian analysis, support of deeper level nodes was not strong, and the bootstrap support for the *Filoviridae*/*Pneumovirinae* clade was only 61.5%. Shimodaira-Hasegawa likelihood ratio testing found that not all trees constrained to monophyly of the *Paramyxoviridae* were significantly worse than the best tree.

4. Discussion

We report the first isolation and partial molecular characterisation of a novel paramyxovirus which we are proposing to name Sunshine virus. This virus is the first paramyxovirus to be isolated in Australia from a non-avian reptile. It is also the first paramyxovirus from a non-avian reptile that does not belong to the genus *Ferlavirus*. The majority of known paramyxoviruses in reptile hosts are found in birds; *Avian metapneumovirus* in the genus *Metapneumovirus* and the entire genus *Avulavirus* use avian hosts. The paramyxoviruses show considerable host diversity by utilising salmonid, reptilian (birds, snakes, lizards and tortoises), and mammalian hosts, and there is no evidence of host-virus codivergence over large-scale paramyxovirus evolution. Ferlaviruses have been found to infect snakes, lizards and tortoises (Marschang et al, 2009). The use of snake hosts by Sunshine virus is likely an independent event from the use of snake hosts by ferlaviruses.

Although Sunshine virus was associated with neurorespiratory disease, a causative association between the presence of Sunshine virus and disease cannot be made from the findings presented here. A transmission study may provide further insight into the pathogenicity of this virus. Despite widespread and repeatable CPE in cell culture, the identity of this virus escaped discovery for over a year while utilising a suite of traditional virological methods. Many methods provided negative or equivocal results (data not shown): transmission electron microscopy (TEM) of infected VH2 cells and supernatant; haemagglutination and haemadsorption assays; neuraminidase activity; effect on viral titre of a DNA synthesis inhibitor (5-bromo-3-deoxyuridine), acid, chloroform and ether. The low viral titre reported in this study ($10^{2.75}$ mL^{-1}) is believed to have contributed to at least some of these equivocal and negative results. Using different cell lines may improve the viral titre. Serum samples from Sunshine virus positive snakes were tested for the presence of anti-ferlaviral antibodies by haemagglutination inhibition (HI) using a neotropical strain of ferlavirus (ATCC VR 1408) as antigen. No HI titre was greater than eight (data not shown) and this was considered to be inconsistent with previous exposure to ferlavirus.

Based on the CPE seen in cell culture, it was suspected that this virus would most likely be a member of one of three families: *Reoviridae*, *Retroviridae* or *Paramyxoviridae*. Due to the difficulty in differentiating endogenous from exogenous retroviruses and the sometimes ambiguous association that retroviruses have with disease, retroviruses were not immediately pursued. Both degenerate and non-degenerate primers for the PCR detection of reoviruses and paramyxoviruses failed to identify this new virus. Of particular note, we were unable to detect Sunshine virus using the primer sets published by Tong et al. (2008). One of these primer sets (PAR-F1, PAR-F2 and PAR-R) was designed to detect all the members of the subfamily *Paramyxovirinae* while another set (PNE-F1, PNE-F2 and PNE-R) was designed to detect all the members of the subfamily *Pneumovirinae*. Tong et al. (2008) were unable to design a set of pan-*Paramyxoviridae* primers and our results suggest that

these primer sets may only be suitable for the detection of novel paramyxoviruses that cluster within either *Paramyxovirinae* or *Pneumovirinae*. In our hands, only high-throughput sequencing using Illumina's® technology was able to identify Sunshine virus as a novel paramyxovirus. Knowledge of Sunshine virus sequence should enable improved future consensus paramyxoviral primer design.

For our phylogenetic investigations, we chose to examine amino acid alignments because of concerns regarding non-lineage factors on viral nucleotide composition bias outweighing the true phylogenetic signal. The paramyxoviruses are divergent to the point that the phosphoproteins and attachment proteins cannot be reliably aligned, and this indicates that the phylogenetic signal from synonymous sites is likely to be significantly weakened by homoplasy. While it has been shown that nucleotide alignments may be moderately more informative than amino acid alignments when looking at vertebrate genes (Townsend et al., 2008), viral evolutionary events such as host switches may cause differential biases in different lineages. In the genus *Atadenovirus*, squamate reptiles appear to be the endemic hosts. These viruses appear to have jumped into birds and mammals in at least two separate events, and in both cases, host jumps were associated with a large AT bias (Wellehan et al., 2004). Experimental cross-species transmission of a feline lentivirus was shown to have a major impact on nucleotide bias (Poss et al., 2006). Host nucleotide composition and host switches also appear to have a significant impact on astrovirus composition (van Hemert et al., 2007).

In all known members of the order *Mononegavirales*, the nucleocapsid gene is near to the 3' end and the polymerase gene is nearest to the 5' end, with the only genes outside of these being NS1 and NS2 before the nucleocapsid in the genus *Pneumovirus*. It is therefore likely that the sequence information generated from this investigation includes all of the coding regions of this new virus. The depth of coverage from the Illumina® High-throughput sequencing is shallow at both the 3' and 5' end of the contig. Therefore, the 17,187 nucleotide contig may not represent the complete genome. 5' or 3' rapid amplification of cDNA ends (RACE) may help in defining terminal ends of this genome. The "rule of six" is the finding that the total number of nucleotides in the viral genome of all members of the subfamily *Paramyxovirinae* is a multiple of six (polyhexameric); this feature is not found in the subfamily *Pneumovirinae* (Kolakofsky et al., 2005; Lamb et al., 2005).The 17,187 nucleotide contig is not divisible by six, but without the certainty that this contig represents the complete genome of Sunshine virus, it cannot be concluded whether this new genome conforms to the "rule of six".

In none of the phylogenetic analyses was there support for the monophyly of the *Paramyxoviridae; Pneumovirinae and Paramyxovirinae did not form a cluster to the exclusion of other virus families*. There was support for monophyly of each paramyxoviral subfamily in all but the matrix analysis, where deep-level resolution was poor. Sunshine virus was the closest relative to *Paramyxovirinae* in all but the matrix analysis, which again lacked resolution. The strongest support for *Paramyxoviridae* monophyly is the presence of a fusion gene, which does not have significant sequence homology with genes of any of the other *Mononegavirales*. However, gene acquisition or loss are also plausible reasons this pattern may be seen. The nucleoprotein analyses identified a clade containing *Rhabdoviridae*/*Paramyxovirinae*/Sunshine virus and a clade containing *Filoviridae*/*Pneumovirinae*,

although likelihood ratio testing did not confirm the significance of these clades. These clades were not seen in the other genes. The two genes with the best resolution, the polymerase and the nucleoprotein, are at opposite ends of the genome. One possible explanation for discordant trees is that these genes do not share the same history and are products of recombination. Although recombination is generally considered uncommon in *Mononegavirales* (McCarthy and Goodman, 2010), it does occur (Chare et al., 2003; Schierup et al., 2005).

We note that another recent phylogenetic analysis of nucleoproteins of the genus *Nyavirus*, divergent members of the *Mononegavirales*, also found support for paraphyly of *Paramyxoviridae* (Mihindukulasuriya et al., 2009). Further understanding of the diversity of the *Mononegavirales* is needed. The small genome size of paramyxoviruses places limitations on phylogenetic resolution, and the best way to improve this is through including further taxa in analyses. The availability of a more complete representation of existing species for comparison results in greater phylogenetic resolution (Flynn et al., 2005; Stefanovic et al., 2004).

Medicine has traditionally waited for viruses to cause epidemics or epizootics before significant surveillance occurs. With increased understanding of virus ecology and evolution, it becomes more feasible to identify probable candidates for future novel disease outbreaks, and increase surveillance. Paramyxoviruses have a clear precedent of jumping hosts and causing significant disease. As examples, *Hendra virus* and *Nipah virus* from bats have recently caused human outbreaks (Field et al., 2007), measles may be derived from rinderpest of cattle (Furuse et al., 2010), and *Human metapneumovirus* is likely of *Avian metapneumovirus* origin (de Graaf et al., 2008). The impact that the discovery of Sunshine virus will have on animal health, including people, by way of broadening the understanding of the paramyxoviruses, is important.

By utilising the Sunshine virus sequence data that was produced by Illumina® high-throughput sequencing, PCR primers have been designed that can be used to detect Sunshine virus in clinical samples (manuscript submitted). Subsequent to this work, Sunshine virus has been sent to one of the co-authors (REM) in Germany for future studies. Initially, this virus will be used for antibody assays in an investigation to see if there is immunological evidence for the presence of this virus in Europe.

Acknowledgements

We would like to thank Dr Mike Bunce for his expertise with the PCRs and high-throughput sequencing.

373

374 # References

375 Adachi, J., Waddell, P.J., Martin, W., Hasegawa, M., 2000. Plastid genome phylogeny and a model of
376 amino acid substitution for proteins encoded by chloroplast DNA. J Mol Evol 50, 348-358.
377 Ahne, W., Batts, W.N., Kurath, G., Winton, J.R., 1999. Comparative Sequence Analyses of Sixteen
378 Reptilian Paramyxoviruses. Virus Res 63, 65-74.
379 Ahne, W., Neubert, W.J., Thomson, I., 1987. Reptilian Viruses: Isolation of Myxovirus-like Particles
380 from the Snake *Elaphe oxycephala*. Zentralbl Veterinarmed B 34, 607-612.
381 Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped
382 BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25,
383 3389-3402.
384 Blahak, S., 1995. Isolation and Characterisation of Paramyxoviruses from Snakes and their
385 Relationship to Avian Paramyxoviruses. Zentralbl Veterinarmed B. 42, 216-224.
386 Chare, E.R., Gould, E.A., Holmes, E.C., 2003. Phylogenetic analysis reveals a low rate of homologous
387 recombination in negative-sense RNA viruses. J Gen Virol 84, 2691-2703.
388 de Graaf, M., Osterhaus, A.D., Fouchier, R.A., Holmes, E.C., 2008. Evolutionary dynamics of human
389 and avian metapneumoviruses. J Gen Virol 89, 2933-2942.
390 Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution
391 39, 783-791.
392 Field, H.E., Mackenzie, J.S., Daszak, P., 2007. Henipaviruses: emerging paramyxoviruses associated
393 with fruit bats. Curr Top Microbiol Immunol 315, 133-159.
394 Flynn, J.J., Finarelli, J.A., Zehr, S., Hsu, J., Nedbal, M.A., 2005. Molecular phylogeny of the carnivora
395 (mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst
396 Biol 54, 317-337.
397 Folsch, D.W., Leloup, P., 1976. [Fatal endemic infection in a serpentarium. Diagnosis, treatment and
398 preventive measures]. Tierarztl Prax 4, 527-536.
399 Franke, J., Essbauer, S., Ahne, W., Blahak, S., 2001. Identification and Molecular Characterisation of
400 18 Paramyxoviruses Isolated from Snakes. Virus Res 80, 67-74.
401 Furuse, Y., Suzuki, A., Oshitani, H., 2010. Origin of measles virus: divergence from rinderpest virus
402 between the 11th and 12th centuries. Virol J 7, 52.
403 Henikoff, S., Henikoff, J.G., 1992. Amino acid substitution matrices from protein blocks. Proceedings
404 of the National Academy of Sciences of the United States of America 89, 10915.
405 ICTV, 2012. Virus Taxonomy: 2011 Release. International Committee on Taxonomy of Viruses.
406 Jacobson, E., Gaskin, J.M., Page, D., Iverson, W.O., Johnson, J.W., 1981. Illness Associated with
407 Paramyxo-like Virus Infection in a Zoologic Collection of Snakes. J Am Vet Med Assoc 179, 1227-
408 1230.
409 Jacobson, E.R., Adams, H.P., Geisbert, T.W., Tucker, S.J., Hall, B.J., Homer, B.L., 1997. Pulmonary
410 Lesions in Experimental Ophidian Paramyxovirus Pneumonia of Aruba Island Rattlesnakes, *Crotalus*
411 *unicolor*. Vet Pathol 34, 450-459.
412 Jacobson, E.R., Gaskin, J.M., Simpson, C.F., Terrell, T.G., 1980. Paramyxo-like Virus Infection in a Rock
413 Rattlesnake. J Am Vet Med Assoc 177, 796-799.
414 Jacobson, E.R., Origgi, F., Pessier, A.P., Lamirande, E.W., Walker, I., Whitaker, B., Stalis, I.H.,
415 Nordhausen, R., Owens, J.W., Nichols, D.K., Heard, D., Homer, B., 2001. Paramyxovirus Infection in
416 Caiman Lizards (*Draecena guianensis*). J Vet Diagn Invest 13, 143-151.
417 Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program.
418 Briefings in Bioinformatics 9, 286.
419 Kolakofsky, D., Roux, L., Garcin, D., Ruigrok, R.W.H., 2005. Paramyxovirus mRNA editing, the 'rule of
420 six'and error catastrophe: a hypothesis. J Gen Virol 86, 1869.

Kolesnikovas, C.K.M., Grego, K.F., Rameh de Albuquerque, L.C., Jacobson, E.R., Monezi, T.A., Mehnert, D.U., Catao-Dias, J.L., 2006. Ophidian Paramyxovirus in Brazilian Vipers (*Bothrops alternatus*). Vet Rec 159, 390-392.

Kurath, G., 2009a. Fish paramyxoviruses: one new species in new genus Aquaparamyxovirus.

Kurath, G., 2009b. Reptile paramyxoviruses: 1 new species and new genus, Ferlavirus.

Kurath, G., Batts, W.N., Ahne, W., Winton, J.R., 2004. Complete Genome Sequence of Fer-de-Lance Virus Reveals a Novel Gene in Reptilian Paramyxoviruses. Journal of Virology 78, 2045-2056.

Lamb, R.A., Collins, P.L., Kolakofsky, D., Melero, J.A., Nagai, Y., Oldstone, M.B.A., Pringle, C.R., Rima, B.K., 2005. Paramyxoviridae, Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, Amsterdam, San Diego, London, pp. 655-668.

Lamb, R.A., Parks, G.D., 2007. Paramyxoviridae: The Viruses and Their Replication, Field's Virology, 5th ed. Lippincott, Williams & Wilkins, a Wolters Kluwer Business, Philadelphia, pp. 1449-1496.

Landolfi, J.A., Terio, K.A., Kinsel, M.J., Langan, J., Zachariah, T.T., Childress, A.L., Wellehan, J.F., Jr., 2010. Orthoreovirus infection and concurrent cryptosporidiosis in rough green snakes (Opheodrys aestivus): pathology and identification of a novel orthoreovirus strain via polymerase chain reaction and sequencing. J Vet Diagn Invest 22, 37-43.

Leary, T.P., Erker, J.C., Chalmers, M.L., Wetzel, J.D., Desai, S.M., Mushahwar, I.K., Dermody, T.S., 2002. Detection of reovirus by reverse transcription-polymerase chain reaction using primers corresponding to conserved regions of the viral L1 genome segment. J Virol Methods 104, 161-165.

Mahy, B.W.J., Kangro, H.O., 1996. Virology Methods Manual. Academic Press, London.

Manvell, R.J., Drury, S.E., Geach, M., Lewis, J.C.M., 2000. Isolation of Ophidian Paramyxovirus type 7 from a Reticulated Python in the UK. Vet Rec 147, 696.

Marschang, R.E., 2011. Viruses Infecting Reptiles. Viruses 3, 2087-2126.

Marschang, R.E., Papp, T., Frost, J.W., 2009. Comparison of paramyxovirus isolates from snakes, lizards and a tortoise. Virus Res 144, 272-279.

McCarthy, A.J., Goodman, S.J., 2010. Reassessing conflicting evolutionary histories of the Paramyxoviridae and the origins of respiroviruses with Bayesian multigene phylogenies. Infect Genet Evol 10, 97-107.

Mihindukulasuriya, K.A., Nguyen, N.L., Wu, G., Huang, H.V., da Rosa, A.P., Popov, V.L., Tesh, R.B., Wang, D., 2009. Nyamanini and midway viruses define a novel taxon of RNA viruses in the order Mononegavirales. J Virol 83, 5109-5116.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees, Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, LA, pp. 1-8.

Myers, T.M., Pieters, A., Moyer, S.A., 1997. A highly conserved region of the Sendai virus nucleocapsid protein contributes to the NP-NP binding domain. Virology 229, 322-335.

Nogueira, M.F., Barrella, T.H., Da Silva, R.J., Lopes, C.A.M., Araujo Junior, J.P., 2002. Isolation of an Ophidian Paramyxovirus (OPMV) in a Captive Rattlesnake (*Crotalus durissus terrificus*) from Botucatu, Sao Paulo State, Brazil. J Venom Anim Toxins 8, 168-173.

Nollens, H.H., Wellehan, J.F., Saliki, J.T., Caseltine, S.L., Jensen, E.D., Van Bonn, W., Venn-Watson, S., 2008. Characterization of a parainfluenza virus isolated from a bottlenose dolphin (Tursiops truncatus). Vet Microbiol 128, 231-242.

Papp, T., Pees, M., Schmidt, V., Marschang, R.E., 2010. RT-PCR diagnosis followed by sequence characterization of paramyxoviruses in clinical samples from snakes reveals concurrent infections within populations and/or individuals. Vet Microbiol 144, 466-472.

Pattengale, N.D., Alipour, M., Bininda-Emonds, O.R., Moret, B.M., Stamatakis, A., 2010. How many bootstrap replicates are necessary? J Comput Biol 17, 337-354.

Poss, M., Ross, H.A., Painter, S.L., Holley, D.C., Terwee, J.A., VandeWoude, S., Rodrigo, A., 2006. Feline lentivirus evolution in cross-species infection reveals extensive G-to-A mutation and selection on key residues in the viral polymerase. J Virol 80, 2728.

Potgieter, L.N.D., Sigler, R.E., Russell, R.G., 1987. Pneumonia in Ottoman Vipers (*Vipera xanthena xanthena*) Associated with a Parainfluenza 2-like Virus. J Wildl Dis 23, 355-360.
Rassa, J.C., Parks, G.D., 1998. Molecular basis for naturally occurring elevated readthrough transcription across the MF junction of the paramyxovirus SV5. Virology 247, 274-286.
Richter, G.A., Homer, B.L., Moyer, S.A., Williams, D.S., Scherba, G., Tucker, S.J., Hall, B.J., Pederson, J.C., Jacobson, E.R., 1996. Characterisation of Paramyxoviruses Isolated from Three Snakes. Virus Res 43, 77-83.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572.
Schierup, M.H., Mordhorst, C.H., Muller, C.P., Christensen, L.S., 2005. Evidence of recombination among early-vaccination era measles virus strains. BMC Evol Biol 5, 52.
Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol 16, 1114-1116.
Stamatakis, A., Hoover, P., Rougemont, J., 2008. A rapid bootstrap algorithm for the RAxML web servers. Syst Biol 57, 758.
Stefanovic, S., Rice, D.W., Palmer, J.D., 2004. Long branch attraction, taxon sampling, and the earliest angiosperms: Amborella or monocots? BMC Evol Biol 4, 35.
Tong, S., Chern, S.W., Li, Y., Pallansch, M.A., Anderson, L.J., 2008. Sensitive and broadly reactive reverse transcription-PCR assays to detect novel paramyxoviruses. J Clin Microbiol 46, 2652-2658.
Townsend, J.P., López-Giráldez, F., Friedman, R., 2008. The phylogenetic informativeness of nucleotide and amino acid sequences for reconstructing the vertebrate tree. J Mol Evol 67, 437-447.
van Hemert, F.J., Berkhout, B., Lukashov, V.V., 2007. Host-related nucleotide composition and codon usage as driving forces in the recent evolution of the Astroviridae. Virology 361, 447-454.
Wellehan, J.F., Johnson, A.J., Harrach, B., Benko, M., Pessier, A.P., Johnson, C.M., Garner, M.M., Childress, A., Jacobson, E.R., 2004. Detection and analysis of six lizard adenoviruses by consensus primer PCR provides further evidence of a reptilian origin for the atadenoviruses. J Virol 78, 13366-13369.
Wellehan, J.F., Jr., Childress, A.L., Marschang, R.E., Johnson, A.J., Lamirande, E.W., Roberts, J.F., Vickers, M.L., Gaskin, J.M., Jacobson, E.R., 2009. Consensus nested PCR amplification and sequencing of diverse reptilian, avian, and mammalian orthoreoviruses. Vet Microbiol 133, 34-42.
Whelan, S., Goldman, N., 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18, 691.
Yong, Y.K., Chong, H.T., Wong, K.T., Tan, C.T., Devi, S., 2008. Aetiology of viral central nervous system infection, a Malaysian study. Neurology Asia 13, 65-71.
Yu, Q., Davis, P., Li, J., Cavanagh, D., 1992. Cloning and sequencing of the matrix protein (M) gene of turkey rhinotracheitis virus reveal a gene order different from that of respiratory syncytial virus. Virology 186, 426-434.
Zangger, N., Muller, M., Pagan, O., 1991. Viral Dermatitis in the Spur-Tailed (*Testudo graeca*) and the Spur-Thighed Tortoise (*Testudo hermanni*) in Switzerland. Proceedings of the 4th International Colloquium on Pathology Medicine of Reptiles and Amphibians, Bad Nauheim, Germany, 17-24.

Figure Legends

Figure 1 Uninfected VH2 cells (above) and VH2 cells infected with Sunshine virus (below). Extensive formation of multinucleate giant cells can be seen in the infected cells. Scale bar represents 100µm.

Figure 2 Map of genomic RNA (3' to 5') of the 17,187nt contig. Numbers represent nucleotide lengths. ORF = open reading frame. Where GenBank similarity could not be validated, putative annotation has been used based on the ORF that is positioned at the corresponding position of other paramyxoviruses.

Figure 3 Bayesian phylogenetic tree of predicted 1,711-2,331 amino acid sequences of mononegaviral RNA-dependent-RNA polymerase based on MAFFT alignment. Bayesian posterior probabilities of clusters as percentages are in bold, and ML bootstrap values for clusters are given to the right. Nyamanini virus (GenBank accession number NC012703) was used as the outgroup. Sunshine virus is bolded. Thick brackets demarcate viral families, medium brackets indicate paramyxoviral subfamilies, and thin brackets demarcate paramyxoviral genera. Sequences retrieved from GenBank include Bornavirus (GenBank accession # NC_001607), Nyamanini (NC_012703), Duvenhage (EU293119), Flanders (AAN73288), Lake Victoria Marburg (YP_001531159), Sudan Ebola (YP_138527), Newcastle disease virus (NC_002617), Avian Paramyxovirus 2 (HM159993), Avian Paramyxovirus 6 (NC_003043), Human Parainfluenzavirus 2 (NC_003443), Simian Parainfluenzavirus 5 (NC_006430), Mapuera (NC_009489), Porcine Rubulavirus (NC_009640), Mumps (NC_002200), Menangle (NC_007620), Tioman (NC_004074), Beilong (NC_007803), J Virus (NC_007454), Canine Distemper Virus (NC_001921), Dolphin Morbillivirus (NC_005283), Measles (NC_001498), Rinderpest (NC_006296), Peste des Petits Ruminants Virus (NC_006383), Mossman (NC_005339), Nariva (FJ362497), Tupaia Paramyxovirus (NC_002199), Hendra (NC_001906), Nipah (NC_002728), Bovine Parainfluenzavirus 3 (NC_002161), Human Parainfluenzavirus 1 (NC_003461), Human Parainfluenzavirus 3 (NC_001796), Sendai (NC_001552), Atlantic Salmon Paramyxovirus (EF646380), Ferlavirus (NC_005084), Avian Metapneumovirus (NC_007652), Human Metapneumovirus (NC_004148), Bovine Respiratory Syncytial Virus (NC_001989), Human Respiratory Syncytial Virus (NC_001781) and Pneumonia Virus of Mice (NC_006579).

Figure 4 Bayesian phylogenetic tree of predicted 529-662 amino acid sequences of paramyxoviral fusion proteins based on MAFFT alignment. Bayesian posterior probabilities of clusters as percentages are in bold, and ML bootstrap values for clusters are given to the right. Atlantic salmon paramyxovirus (EF646380) was used as the outgroup. Sunshine virus is bolded. Medium brackets indicate paramyxoviral subfamilies, and thin brackets demarcate paramyxoviral genera. Sequences retrieved from GenBank are from the same genomes as the L genes used in Figure 3.

Figure 5 Bayesian phylogenetic tree of predicted 142-377 amino acid sequences of mononegaviral matrix proteins based on MAFFT alignment. Multifurcations are marked with arcs. Bayesian posterior probabilities of clusters as percentages are in bold, and ML bootstrap values for clusters are given to the right. Bornavirus (NC_001607) was used as the outgroup. Sunshine virus is bolded. Thick brackets demarcate viral families, medium brackets indicate paramyxoviral subfamilies, and thin brackets demarcate paramyxoviral genera. Sequences retrieved from GenBank are from the same genomes as the L genes used in Figure 3.

Figure 6 Bayesian phylogenetic tree of predicted 370-738 amino acid sequences of mononegaviral nucleoproteins based on MAFFT alignment. Bayesian posterior probabilities of clusters as percentages are in bold, and ML bootstrap values for clusters are given to the right. Nyamanini virus

559 (GenBank accession number NC012703) was used as the outgroup. Sunshine virus is bolded. Thick
560 brackets demarcate viral families, medium brackets indicate paramyxoviral subfamilies, and thin
561 brackets demarcate paramyxoviral genera. Sequences retrieved from GenBank are from the same
562 genomes as the L genes used in Figure 3.

Tables

Primer set (number of nucleotides from 3' end of polymerase gene)	**PCR product size (bp)**
SunshineS1 (2444): 5'GGAAAGGGAGGTCTATG SunshineAS1 (2596): 5'ATTCAACATCTGGGGTC	153
SunshineS2 (2240): 5'TTCAAGGAGATAACCAGG SunshineAS2 (2469): 5'CGGGATTCCCATAGAC	230
SunshineS2 (2240): 5'TTCAAGGAGATAACCAGG SunshineAS1 (2596): 5'ATTCAACATCTGGGGTC	357

Table 1. Primer sequences, and anticipated amplicon size, used for the detection of the polymerase gene of Sunshine virus. PCR = polymerase chain reaction. bp = base pairs.

Gene	**Start**	**Position**	**End**	**Position**	**Intergenic Space**
Nucleoprotein	**CUCUCUCU**	32	**AUUU**UG**UUUUUUU**	1,570	0 nucleotides
Unknown	**CUCUCUCU**	1,583	**GUU**GU**AUUUUU**	2,391	32 nucleotides
Phosphoprotein	**CUCUCUCU**	2,434	**G**G**UUAAUUUUUU**	4,328	0 nucleotides
Matrix	**CUCUCUCUCU**	4,340	C**UU**CUG**UUUU**	5,496?	?
			UA**UU**GC**UUUU**	5,520?	?
			UA**UUAAUUUUUUU**	5,708?	54 nucleotides
Fusion	U**U**U**UCU**G**U**A**U**	5,513?			
	U**U**U**UCU**U**U**	5,580?			
	G**UCU**U**U**A**U**	5,589?			
	CUCUCUCU	5,776?	**G**G**UUAAUUUUUU**	7,666	10 nucleotides
Attachment	**CUCUCUCU**	7,688	**GUUUAAUUUU**	10,187	31 nucleotides
			AUUUAAUUUUUUU	10,208	7 nucleotides
Polymerase	**CUCUC**A**CU**	10,228	**AUUUAAUUUUUU**	17,044	
Consensus	**CUCUCuCU**		**RuUuaa(U)**$_{4-8}$		

Table 2. Conserved motifs in Sunshine virus consistent with paramyxoviral gene transcription start and end signals. Sequences are given 3'-5' in genomic antisense. Positions are nucleotide distances from the 3' end of the obtained sequence. Sites matching the consensus motifs are in bold.

Appendix

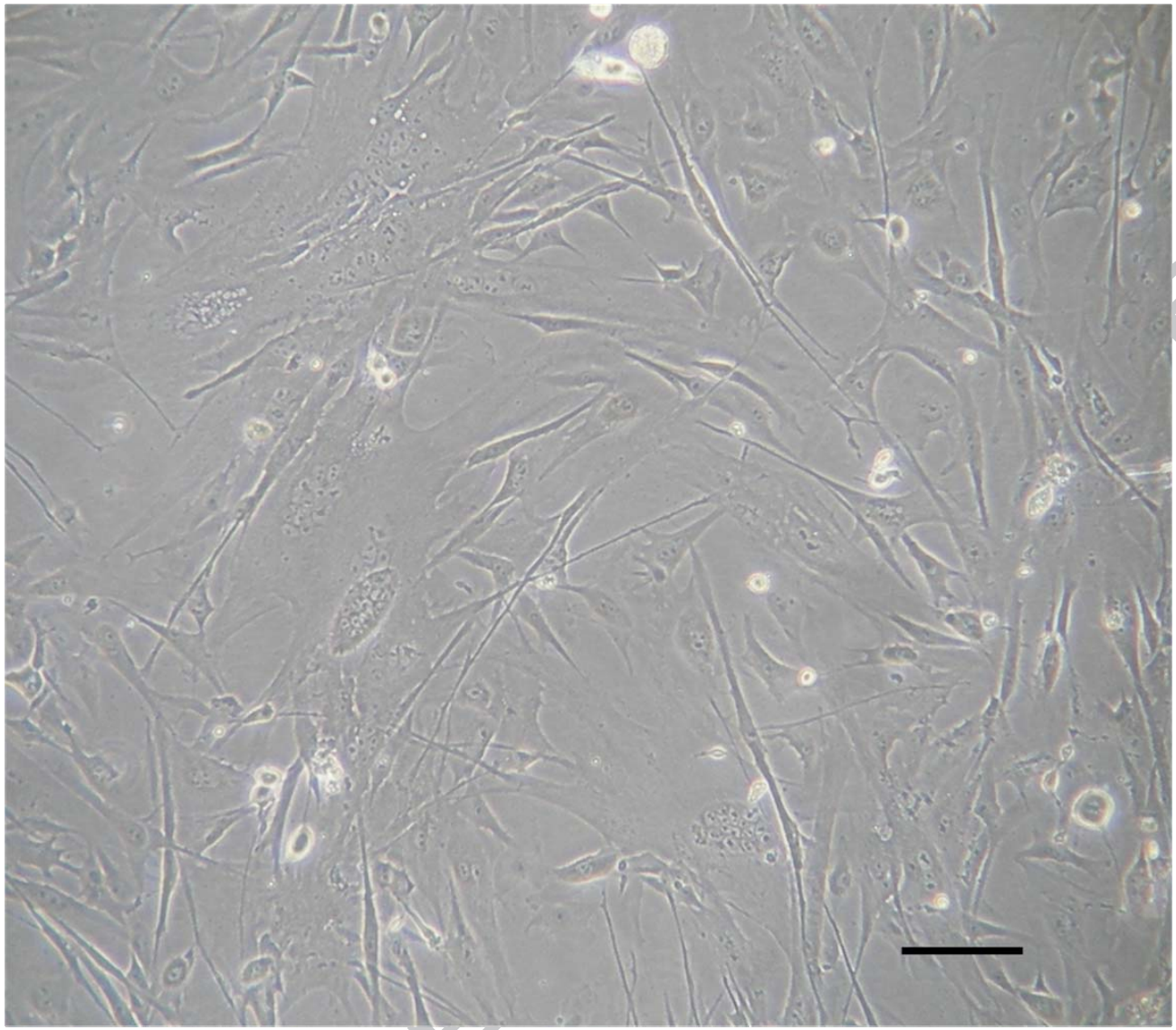
Virus	Genome region	Primers (5' → 3'), amplicon size		Reference
		First round	Second round	
Paramyxoviridae	Polymerase (L)	qS2 (GTTATGGCAAATCATGCTGCGATACCTTA) qAS2 (CTGATGGGAGATAATGCCTTGTCCTTCAT), **157bp**	Single round only	Designed in-house
		DegenParamyxoS (GGIGGKATWGAAGGWTWITGYCAAAAAMTRTGGAC) DegenParamyxoAS (TKAYTGCWATTGMTTGATTGTCWCC), **109bp**	Single round only	Designed in-house
		L5 (GCAGAGATTTTCTCTTTCTT) L6 (AGCTCTCATTTTGTATGTCAT), **627bp**	L7 (TAGAGGCTGTTACTGCTGC) L8 (CATCTTGGCAAATAATCTGCC), **566bp**	(Ahne et al., 1999)
		PAR-F1 (GAAGGITATTGTCAIAARNTNTGGAC) PAR-R (GCTGAAGTTACIGGITCICCDATRTTNC), **662bp***	PAR-F2 (GTTGCTTCAATGGTTCARGGNGAYAA) PAR-R, **584bp***	(Tong et al., 2008)
		PAR-F2 PAR-R, **584bp***	Single round only	(Tong et al., 2008)
		PNE-F1 (GTGTAGGTAGIATGTTYGCNATGCARCC) PNE-R (GTCCCACAAITTTTGRCACCANCCYTC), **488bp***	PNE-F2 (ACTGATCTIAGYAARTTYAAYCARGC) PNE-R, **264bp***	(Tong et al., 2008)
		PNE-F2 PNE-R, **264bp***	Single round only	(Tong et al., 2008)
		pmxF (TACTGCCTNAATTGGAGATATGA) pmxR (CCTTCTATACCCCCTCTAGGATA), **224bp***	Single round only	(Yong et al., 2008)
		PMV14610F (TTTGCNAAAATGACNTACAAAATGAG) PMV15353R (GAAGGNTATTGNCAGAAGNTATGGAC), **744bp**	Single round only	(Nollens et al., 2008)
Reoviridae	Polymerase (L1)	ReoS (AACAACAGCACGATGATGAA) ReoAS1 (ACCATTATCCCATCATCACC), **137bp**	Single round only	Designed in-house
		L1.rv5 (GCATCCATTGTAAATGACGAGTCTG) L1.rv6 (CTTGAGATTAGCTCTAGCATCTTCTG), **416bp**	L1.rv7 (GCTAGGCCGATATCGGGAATGCAG) L1.rv8 (GTCTCACTATTCACCTTACCAGCAG), **344bp**	(Leary et al., 2002)
		L1.rv5m (CTGCATCCATTGTAAATGACGAGTC) L1.rv4m (GCTATGTCATATTTCCATCCGAATTC), **455bp**	L1.rv6 L1.rv7, **374bp**	(Leary et al., 2002)
		1607F (CARMGNCGNSCHMGHTCHATHATGCC) 2608R (TAVAYRAAVGWCCASMHNGGRTAYTG), **1,053bp**	2090F (GGBTCMACNGCYACYTCBACYGAGCA) 2334R (CDATGTCRTAHWYCCANCCRAA), **292bp**	(Wellehan et al., 2009)
		2090F 2334R, **292bp**	2090F 2200R (CCRTCRTCWCCYTGRCAKAYRTARTT), **162bp**	(Landolfi et al., 2010)

Table A.1. Primers used for the detection of the polymerase gene (L) of paramyxoviruses, especially ferlaviruses. R = A or G; Y = C or T; W = A or T; K = G or T; S = C or G; H = A, C or T; M = A or C; D = A, G or T; V = A, G or C; B = G, C or T; N = A, C, G or T, I=inosine. *Information not provided in original manuscript but estimated from binding sites in other paramyxoviruses.

Highlights

- A novel paramyxovirus was isolated from Australian pythons
- High throughput sequencing was used to identify this virus
- Phylogeny shows that this virus clusters outside of both paramyxoviral subfamilies
- The name Sunshine virus is proposed for this new virus

Fig. 1



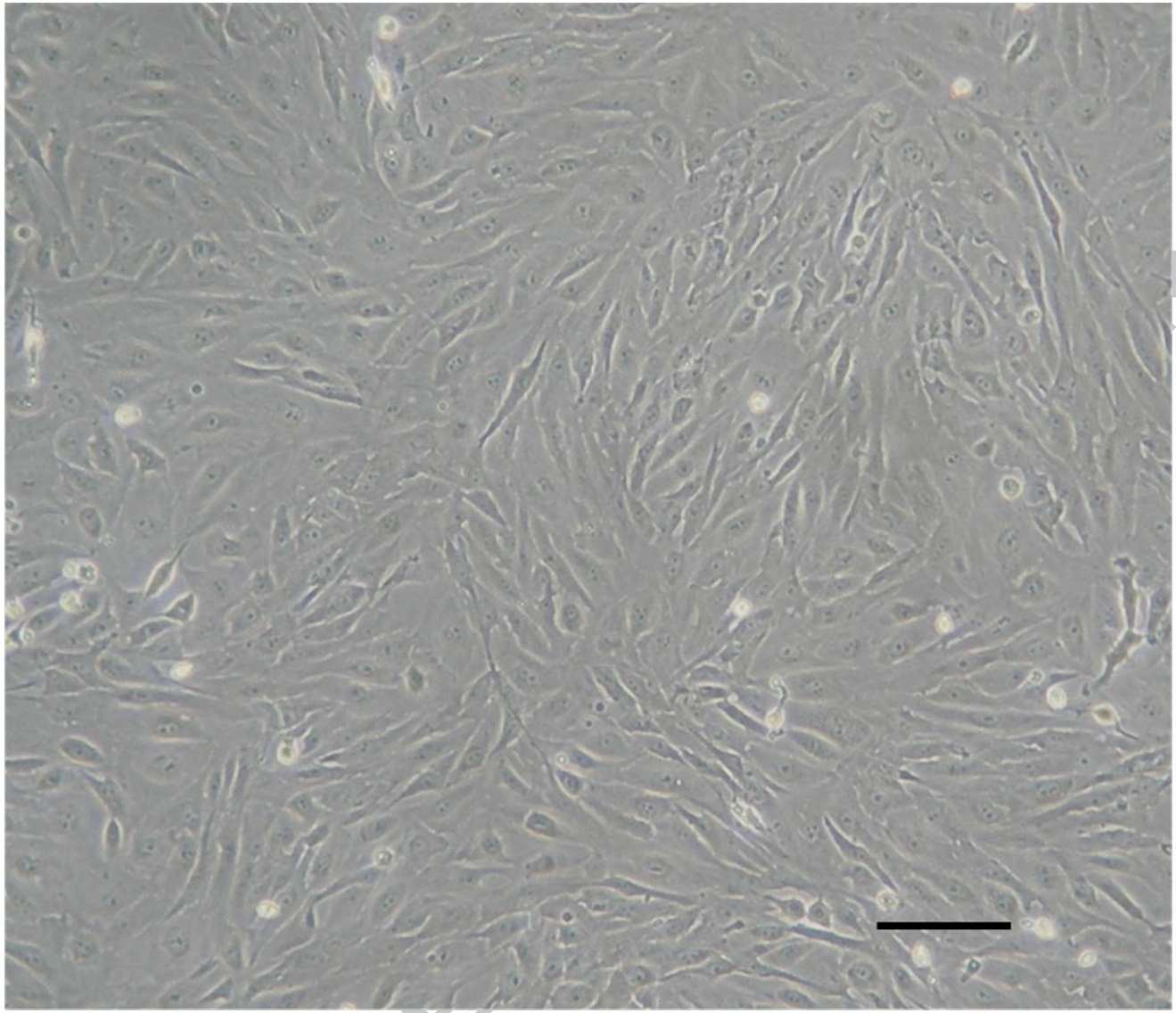
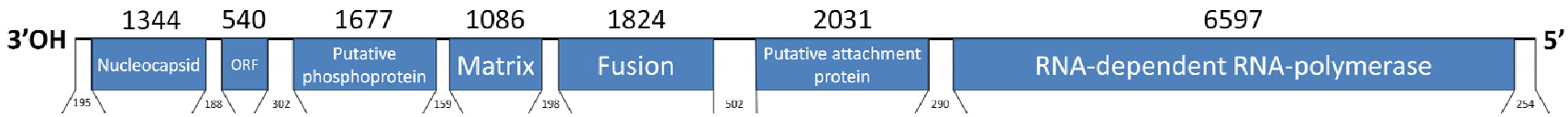
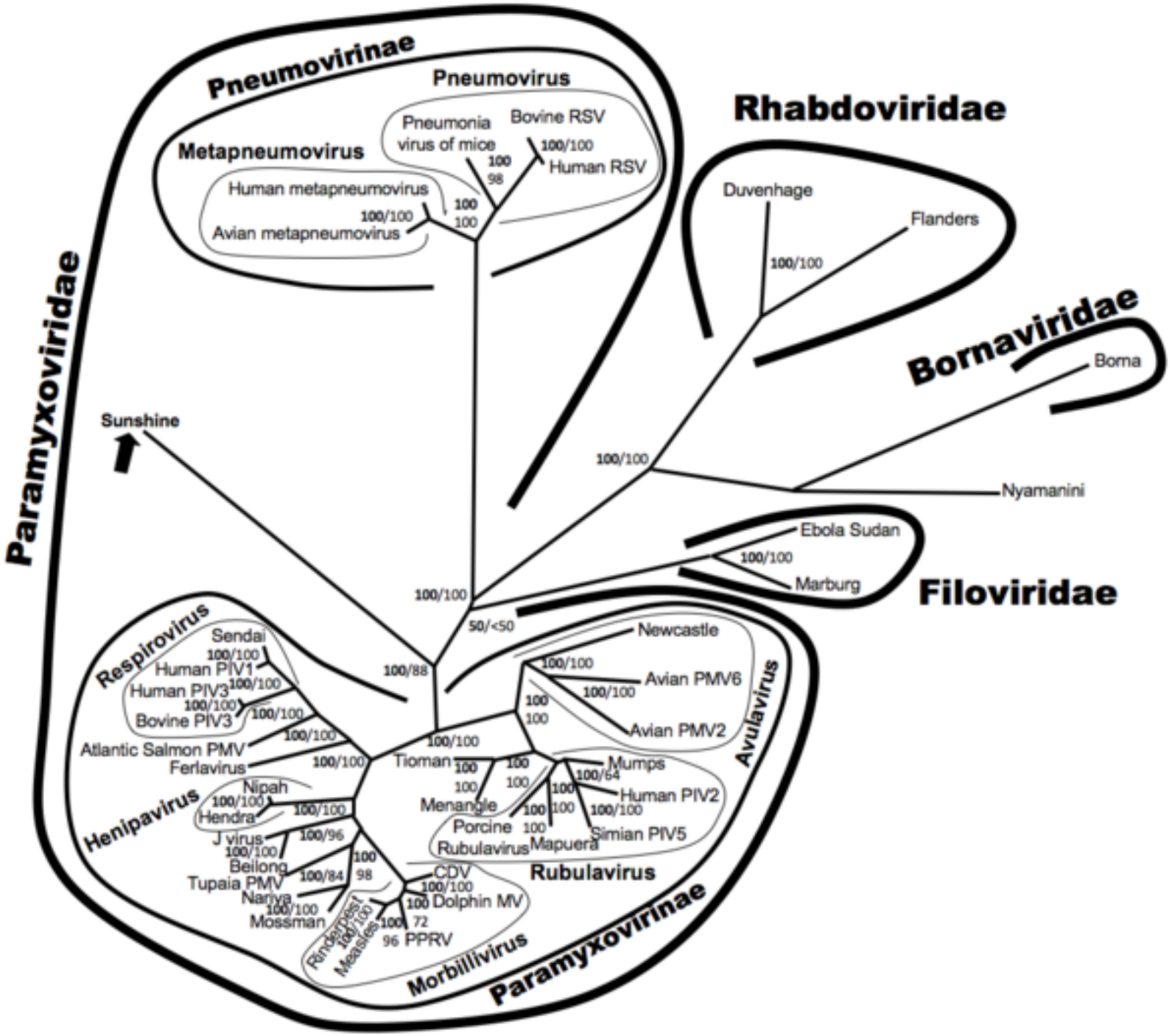
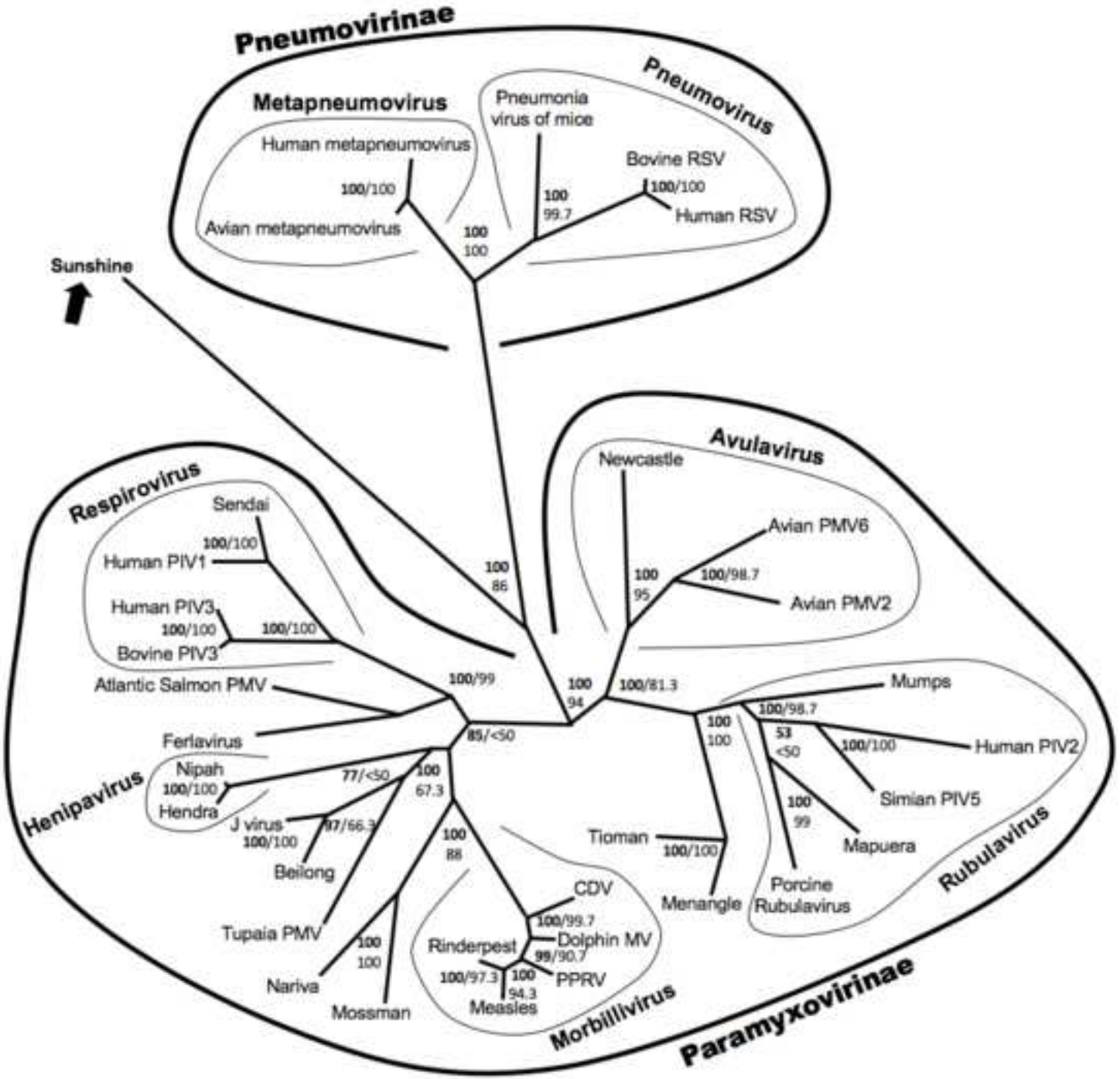


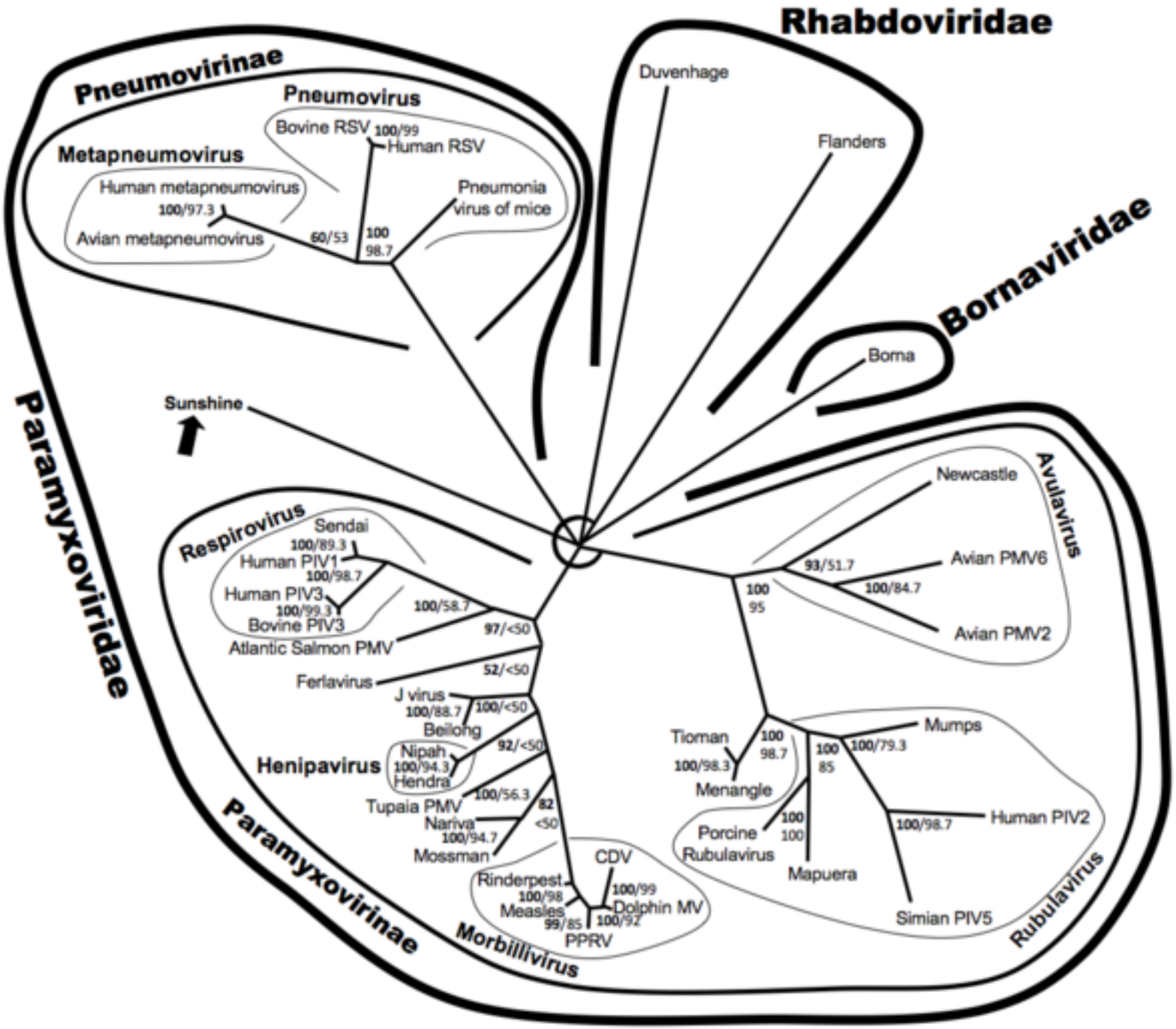
Fig. 2



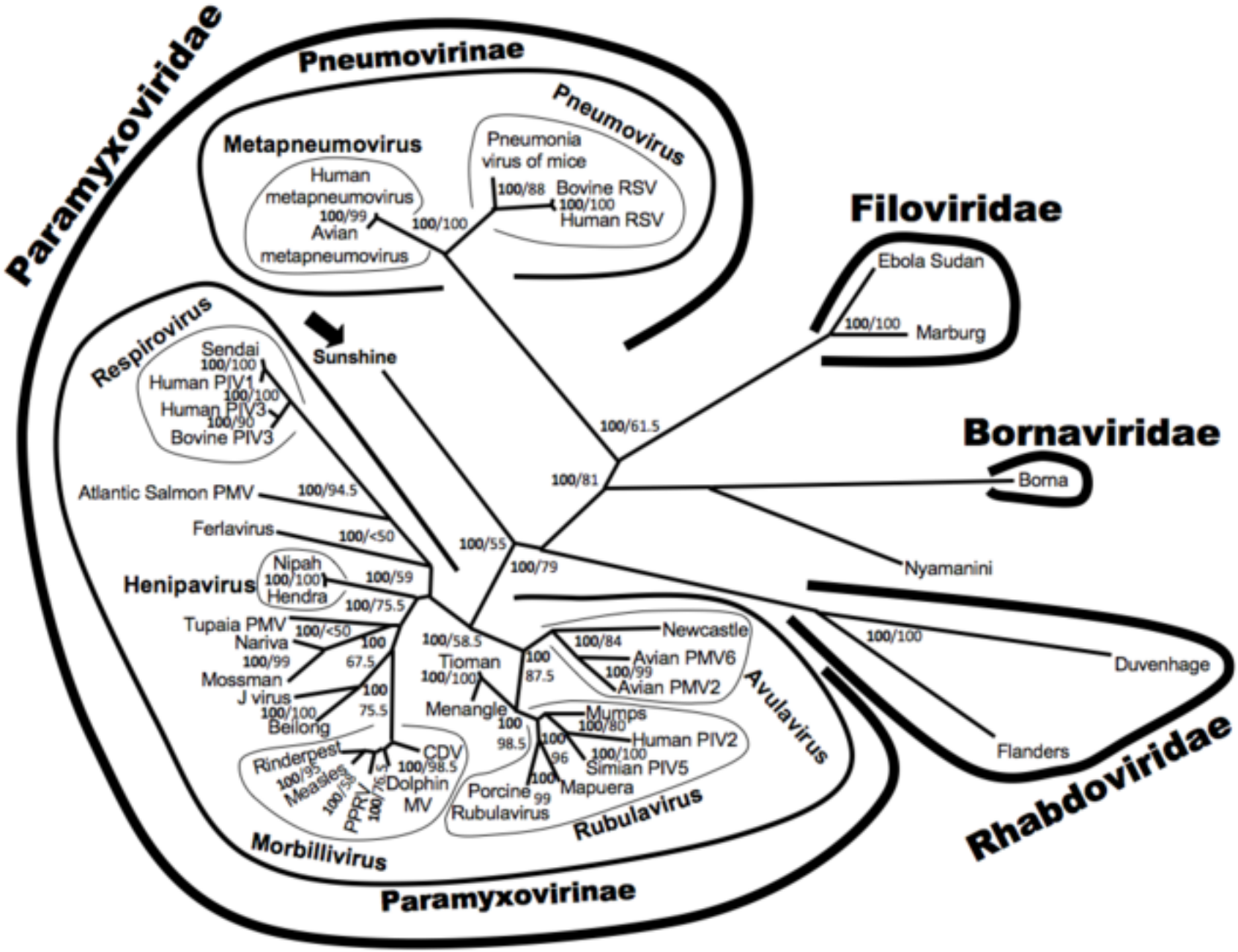




F



M



N